BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: June 18, 2003, 17:14:36 ; Search time 77.2272 Seconds
(without alignments)
504.414 Million cell updates/sec

Title: US-09-807-933B-5
Perfect score: 1956
Sequence: 1 MKFLTIASSAILALAVGTEM.....TYKQVTCPKAITAKSGCSRK 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	747	38.2	225	9	US-10-007-521-2	Sequence 2, Appli
2	747	38.2	297	9	US-10-007-521-4	Sequence 4, Appli
3	747	38.2	308	9	US-10-007-521-6	Sequence 6, Appli
4	745	38.1	299	9	US-10-007-521-12	Sequence 12, Appl
5	735	37.6	205	9	US-09-261-329-7	Sequence 7, Appli
6	732	37.4	201	9	US-09-261-329-5	Sequence 5, Appli
7	724	37.0	201	9	US-09-261-329-4	Sequence 4, Appli
8	722.5	36.9	376	10	US-09-735-787-4	Sequence 4, Appli
9	713	36.5	203	9	US-09-261-329-6	Sequence 6, Appli
10	710.5	36.3	349	9	US-10-007-521-10	Sequence 10, Appl
11	698.5	35.7	310	9	US-10-007-521-22	Sequence 22, Appl
12	689.5	35.3	202	9	US-09-261-329-3	Sequence 3, Appli
13	687	35.1	203	9	US-09-261-329-9	Sequence 9, Appli
14	687	35.1	222	9	US-10-007-521-14	Sequence 14, Appl
15	687	35.1	294	9	US-10-007-521-24	Sequence 24, Appl
16	676.5	34.6	305	10	US-09-735-787-2	Sequence 2, Appli
17	675.5	34.5	202	9	US-09-261-329-1	Sequence 1, Appli
18	674	34.5	235	1	US-08-841-636A-31	Sequence 31, Appl
19	660	33.7	226	9	US-10-007-521-16	Sequence 16, Appl
20	660	33.7	293	9	US-10-007-521-20	Sequence 20, Appl
21	660	33.7	298	9	US-10-007-521-18	Sequence 18, Appl
22	656.5	33.6	295	9	US-10-007-521-8	Sequence 8, Appli
23	652	33.3	203	9	US-09-261-329-8	Sequence 8, Appli
24	647.5	33.1	202	9	US-09-261-329-2	Sequence 2, Appli
25	485	24.8	211	9	US-09-261-329-11	Sequence 11, Appl
26	458	23.4	235	9	US-09-261-329-10	Sequence 10, Appl
27	436	22.3	138	9	US-10-007-521-26	Sequence 26, Appl
28	219.5	11.2	75	9	US-10-007-521-32	Sequence 32, Appl
29	211	10.8	2916	9	US-10-123-155-69	Sequence 69, Appl
30	210	10.7	1660	9	US-10-184-644-147	Sequence 147, App
31	210	10.7	1660	9	US-10-184-634-147	Sequence 147, App
32	206.5	10.6	3127	9	US-10-184-644-83	Sequence 83, Appl
33	206.5	10.6	3127	9	US-10-184-634-83	Sequence 83, Appl
34	205	10.5	1591	9	US-10-073-912-13	Sequence 13, Appl
35	202.5	10.4	18636	9	US-10-073-912-17	Sequence 17, Appl
36	201	10.3	1621	9	US-10-123-155-415	Sequence 415, App
37	201	10.3	2014	9	US-10-123-155-51	Sequence 51, Appl
38	200.5	10.3	1819	9	US-10-184-644-39	Sequence 39, Appl
39	200.5	10.3	1819	9	US-10-184-634-39	Sequence 39, Appl
40	200	10.2	2014	9	US-10-123-155-547	Sequence 547, App
41	198.5	10.1	475	10	US-09-826-752-14	Sequence 14, Appl
42	198.5	10.1	2698	9	US-10-123-155-25	Sequence 25, Appl
43	198.5	10.1	3401	9	US-10-184-644-411	Sequence 411, App
44	198.5	10.1	3401	9	US-10-184-634-411	Sequence 411, App
45	197	10.1	1915	9	US-10-184-644-275	Sequence 275, App

ALIGNMENTS

```
RESULT 1
US-10-007-521-2
; Sequence 2, Application US/10007521
; Publication No. US20030054539A1
;  GENERAL INFORMATION:
;       APPLICANT: Schulein, Martin
;                  Andersen, Lene N.
;                  Lassen, Soren F.
;                  Kauppinen, Markus S.
;                  Lange, Lene
;                  Nielsen, Ruby I.
;                  Ihara, Michiko
;                  Takagi, Shinobu
;       TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
;       NUMBER OF SEQUENCES: 109
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US200300545
;            STREET: 405 Lexington Avenue, 64th Floor
;            CITY: New York
;            STATE: New York
;            COUNTRY: United States of America
;            ZIP: 10174-6401
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/007,521
;            FILING DATE: 10-Dec-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/651,136
;            FILING DATE: 21-MAY-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Lambiris, Elias J.
;            REGISTRATION NUMBER: 33,728
;            REFERENCE/DOCKET NUMBER: 4366.200-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 212-867-0123
;            TELEFAX: 212-878-9655
;  INFORMATION FOR SEQ ID NO: 2:
```

```
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 225 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-007-521-2

  Query Match             38.2%;  Score 747;  DB 9;  Length 225;
  Best Local Similarity   57.5%;  Pred. No. 3.8e-45;
  Matches  130;  Conservative   38;  Mismatches   48;  Indels   10;  Gaps    5;

Qy        140 SASTNYSAVSGGA----SGNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLAD-N 194
              ||:| : |:   |    || |:|||||||||||||:||||   :||| :|:|:   | |
Db          4 SATTGFLALPVLALDQLSGIGQTTRYWDCCKPSCAWPGKGP-SSPVQACDKNDNPLNDGG 62

Qy        195 NTQNGC-VGGSSYTCNDNQPWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKG 253
              :|::||  |||:| |:   || |||:|:||:||  ::| ||: |||||:||||||  | |
Db         63 STRSGCDAGGSAYMCSSQSPWAVSDELSYGWAAVKLAGSSESQWCCACYELTFTSGPVAG 122

Qy        254 KKMVVQVTNTGSDLGSNTGAHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASD 313
              |||:|| |||| ||| |   |||| :|||||||:| |  |:||| :||| ||||: |  :
Db        123 KKMIVQATNTGGDLGDN---HFDLAIPGGGVGIFNACTDQYGAPPNGWGDRYGGIHSKEE 179

Qy        314 CSNLPSALQAGCKWRFGWFKNADNPTMTYKQVTCPKAITAKSGCSR 359
              | : | ||: || ||| ||:|||||::|:::| ||  :|:||||||
Db        180 CESFPEALKPGCNWRFDWFQNADNPSVTFQEVACPSELTSKSGCSR 225


RESULT 2
US-10-007-521-4
; Sequence 4, Application US/10007521
; Publication No. US20030054539A1
;   GENERAL INFORMATION:
;        APPLICANT: Schulein, Martin
;                   Andersen, Lene N.
;                   Lassen, Soren F.
;                   Kauppinen, Markus S.
;                   Lange, Lene
;                   Nielsen, Ruby I.
;                   Ihara, Michiko
;                   Takagi, Shinobu
;        TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
;        NUMBER OF SEQUENCES: 109
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US200300545
;             STREET: 405 Lexington Avenue, 64th Floor
;             CITY: New York
;             STATE: New York
;             COUNTRY: United States of America
;             ZIP: 10174-6401
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/007,521
;             FILING DATE: 10-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/651,136
;             FILING DATE: 21-MAY-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lambiris, Elias J.
;             REGISTRATION NUMBER: 33,728
;             REFERENCE/DOCKET NUMBER: 4366.200-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 212-867-0123
;             TELEFAX: 212-878-9655
;   INFORMATION FOR SEQ ID NO: 4:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 297 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-007-521-4

  Query Match             38.2%;  Score 747;  DB 9;  Length 297;
  Best Local Similarity   57.5%;  Pred. No. 5.1e-45;
  Matches  130;  Conservative   38;  Mismatches   48;  Indels   10;  Gaps    5;

Qy        140 SASTNYSAVSGGA----SGNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLAD-N 194
              ||:| : |:   |    || |:|||||||||||||:||||   :||| :|:|:   | |
Db          4 SATTGFLALPVLALDQLSGIGQTTRYWDCCKPSCAWPGKGP-SSPVQACDKNDNPLNDGG 62

Qy        195 NTQNGC-VGGSSYTCNDNQPWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKG 253
              :|::||  |||:| |:   || |||:|:||:||  ::| ||: |||||:||||||  | |
Db         63 STRSGCDAGGSAYMCSSQSPWAVSDELSYGWAAVKLAGSSESQWCCACYELTFTSGPVAG 122

Qy        254 KKMVVQVTNTGSDLGSNTGAHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASD 313
              |||:|| |||| ||| |   |||| :|||||||:| |  |:||| :||| ||||: |  :
Db        123 KKMIVQATNTGGDLGDN---HFDLAIPGGGVGIFNACTDQYGAPPNGWGDRYGGIHSKEE 179

Qy        314 CSNLPSALQAGCKWRFGWFKNADNPTMTYKQVTCPKAITAKSGCSR 359
              | : | ||: || ||| ||:|||||::|:::| ||  :|:||||||
Db        180 CESFPEALKPGCNWRFDWFQNADNPSVTFQEVACPSELTSKSGCSR 225


RESULT 3
US-10-007-521-6
; Sequence 6, Application US/10007521
; Publication No. US20030054539A1
;   GENERAL INFORMATION:
;        APPLICANT: Schulein, Martin
;                   Andersen, Lene N.
;                   Lassen, Soren F.
;                   Kauppinen, Markus S.
;                   Lange, Lene
;                   Nielsen, Ruby I.
;                   Ihara, Michiko
;                   Takagi, Shinobu
;        TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
;        NUMBER OF SEQUENCES: 109
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US200300545
;             STREET: 405 Lexington Avenue, 64th Floor
;             CITY: New York
;             STATE: New York
;             COUNTRY: United States of America
;             ZIP: 10174-6401
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/007,521
;             FILING DATE: 10-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/651,136
;             FILING DATE: 21-MAY-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lambiris, Elias J.
;             REGISTRATION NUMBER: 33,728
;             REFERENCE/DOCKET NUMBER: 4366.200-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 212-867-0123
;             TELEFAX: 212-878-9655
;   INFORMATION FOR SEQ ID NO: 6:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 308 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
```

```
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-007-521-6

  Query Match           38.2%;  Score 747;  DB 9;  Length 308;
  Best Local Similarity   57.5%;  Pred. No. 5.3e-45;
  Matches  130;  Conservative   38;  Mismatches   48;  Indels   10;  Gaps    5;

Qy        140 SASTNYSAVSGGA----SGNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLAD-N 194
              ||:| : |:   |    || |:|||||||||||:||||   :||| :|:|:   | |
Db          4 SATTGFLALPVLALDQLSGIGQTTRYWDCCKPSCAWPGKGP-SSPVQACDKNDNPLNDGG 62

Qy        195 NTQNGC-VGGSSYTCNDNQPWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKG 253
              :|::||  |||:| |:   || |||:|:||:||  ::| ||: |||||:||||||  | |
Db         63 STRSGCDAGGSAYMCSSQSPWAVSDELSYGWAAVKLAGSSESQWCCACYELTFTSGPVAG 122

Qy        254 KKMVVQVTNTGSDLGSNTGAHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASD 313
              |||:|| |||| ||| |   |||| :|||||||:| |  |:||| :||| ||||: |  :
Db        123 KKMIVQATNTGGDLGDN---HFDLAIPGGGVGIFNACTDQYGAPPNGWGDRYGGIHSKEE 179

Qy        314 CSNLPSALQAGCKWRFGWFKNADNPTMTYKQVTCPKAITAKSGCSR 359
              | : | ||: || ||| ||:|||||::|:::| ||  :|:||||||
Db        180 CESFPEALKPGCNWRFDWFQNADNPSVTFQEVACPSELTSKSGCSR 225


RESULT 4
US-10-007-521-12
; Sequence 12, Application US/10007521
; Publication No. US20030054539A1
;   GENERAL INFORMATION:
;       APPLICANT: Schulein, Martin
;                  Andersen, Lene N.
;                  Lassen, Soren F.
;                  Kauppinen, Markus S.
;                  Lange, Lene
;                  Nielsen, Ruby I.
;                  Ihara, Michiko
;                  Takagi, Shinobu
;       TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
;       NUMBER OF SEQUENCES: 109
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US200300545
;            STREET: 405 Lexington Avenue, 64th Floor
;            CITY: New York
;            STATE: New York
;            COUNTRY: United States of America
;            ZIP: 10174-6401
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/007,521
;            FILING DATE: 10-Dec-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/651,136
;            FILING DATE: 21-MAY-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Lambiris, Elias J.
;            REGISTRATION NUMBER: 33,728
;            REFERENCE/DOCKET NUMBER: 4366.200-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 212-867-0123
;            TELEFAX: 212-878-9655
;   INFORMATION FOR SEQ ID NO: 12:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 299 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-007-521-12

  Query Match           38.1%;  Score 745;  DB 9;  Length 299;
  Best Local Similarity   60.4%;  Pred. No. 7.1e-45;
  Matches  128;  Conservative   36;  Mismatches   42;  Indels    6;  Gaps    2;

Qy        148 VSGGASGNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLADNNTQNGCVGGSSYT 207
              |:  |||:|::|||||||||||:|||||  |: || :|: : : |:| | |:|| |||:|:
Db         18 VASAASGSGQSTRYWDCCKPSCAWPGKAAVSQPVYACDANFQRLSDFNVQSGCNGGSAYS 77

Qy        208 CNDNQPWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTGSDL 267
              | |  || |:|:|||||||  ||:||||::|||||: |||||  | || |||| |:|| ||
Db         78 CADQTPWAVNDNLAYGFAATSIAGGSESSWCCACYALTFTSGPVAGKTMVVQSTSTGGDL 137

Qy        268 GSNTGAHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKW 327
              |||    ||: |||||||||:|||::|:|      ||:|||:||   | : |: |: ||:|
Db        138 GSN---QFDIAMPGGGVGIFNGCSSQFGGLP---GAQYGGISSRDQCDSFPAPLKPGCQW 191

Qy        328 RFGWFKNADNPTMTYKQVTCPKAITAKSGCSR 359
              || ||:|||||| |::|| ||  | |:||| |
Db        192 RFDWFQNADNPTFTFQQVQCPAEIVARSGCKR 223


RESULT 5
US-09-261-329-7
; Sequence 7, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Kim
;  APPLICANT: Schulein, Martin
;  APPLICANT: Christiansen, Lars
;  APPLICANT: Damgaard, Bo
;  APPLICANT: Von Der Osten,  Claus
;  TITLE OF INVENTION: Cellulase Variants
;  FILE REFERENCE: 4887.204-US
;  CURRENT APPLICATION NUMBER: US/09/261,329
;  CURRENT FILING DATE: 1999-03-03
;  EARLIER APPLICATION NUMBER: 1013/96
;  EARLIER FILING DATE: 1996-09-17
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
;   LENGTH: 205
;   TYPE: PRT
;   ORGANISM: Cellulase variants
US-09-261-329-7

  Query Match           37.6%;  Score 735;  DB 9;  Length 205;
  Best Local Similarity   59.1%;  Pred. No. 2.4e-44;
  Matches  123;  Conservative   35;  Mismatches   44;  Indels    6;  Gaps    4;

Qy        154 GNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLAD-NNTQNGC-VGGSSYTCNDN 211
              | |:|||||||||||:||||   :||| :|:|:     |  :|::||  |||:| |:
Db          1 GIGQTTRYWDCCKPSCAWPGKGP-SSPVQACDKNDNPFNDGGSTRSGCDAGGSAYMCSSQ 59

Qy        212 QPWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTGSDLGSNT 271
              || |||:|:||:||  ::| ||: |||||:||||||  | ||||:|| |||| ||| |
Db         60 SPWAVSDELSYGWAAVKLAGSSESQWCCACYELTFTSGPVAGKKMIVQATNTGGDLGDN- 118

Qy        272 GAHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKWRFGW 331
              |||| :|||||||:| |  |:||| :||| ||||: |  :| : | ||: || ||| |
Db        119 --HFDLAIPGGGVGIFNACTDQYGAPPNGWGDRYGGIHSKEECESFPEALKPGCNWRFDW 176

Qy        332 FKNADNPTMTYKQVTCPKAITAKSGCSR 359
              |:|||||::|:::| ||  :|:||||||
Db        177 FQNADNPSVTFQEVACPSELTSKSGCSR 204


RESULT 6
US-09-261-329-5
; Sequence 5, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Andersen, Kim
;  APPLICANT: Schulein, Martin
;  APPLICANT: Christiansen, Lars
;  APPLICANT: Damgaard, Bo
;  APPLICANT: Von Der Osten,  Claus
;  TITLE OF INVENTION: Cellulase Variants
;  FILE REFERENCE: 4887.204-US
;  CURRENT APPLICATION NUMBER: US/09/261,329
;  CURRENT FILING DATE: 1999-03-03
;  EARLIER APPLICATION NUMBER: 1013/96
;  EARLIER FILING DATE: 1996-09-17
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 201
;   TYPE: PRT
;   ORGANISM: Cellulase variants
US-09-261-329-5

  Query Match             37.4%;  Score 732;  DB 9;  Length 201;
  Best Local Similarity   60.7%;  Pred. No. 3.8e-44;
  Matches  125;  Conservative   35;  Mismatches   40;  Indels    6;  Gaps    2;

Qy        154 GNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLADNNTQNGCVGGSSYTCNDNQP 213
              |:|::|||||||||||:|||| |: || :|: : : |:| | |:|| |||:|:| | |
Db          1 GSGQSTRYWDCCKPSCAWPGKAAVSQPVYACDANFQRLSDFNVQSGCNGGSAYSCADQTP 60

Qy        214 WVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTGSDLGSNTGA 273
              | |:|:|||||| ||:||||::|||||: ||||| | || |||| |:|| |||||
Db         61 WAVNDNLAYGFAATSIAGGSESSWCCACYALTFTSGPVAGKTMVVQSTSTGGDLGSN--- 117

Qy        274 HFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKWRFGWFK 333
              ||: |||||||:|||::|:|       ||:|||:||   | : |: |: ||:||| ||:
Db        118 QFDIAMPGGGVGIFNGCSSQFGGLP---GAQYGGISSRDQCDSFPAPLKPGCQWRFDWFQ 174

Qy        334 NADNPTMTYKQVTCPKAITAKSGCSR 359
              |||||| |::|| ||  | |:||| |
Db        175 NADNPTFTFQQVQCPAEIVARSGCKR 200


RESULT 7
US-09-261-329-4
; Sequence 4, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Kim
;  APPLICANT: Schulein, Martin
;  APPLICANT: Christiansen, Lars
;  APPLICANT: Damgaard, Bo
;  APPLICANT: Von Der Osten,  Claus
;  TITLE OF INVENTION: Cellulase Variants
;  FILE REFERENCE: 4887.204-US
;  CURRENT APPLICATION NUMBER: US/09/261,329
;  CURRENT FILING DATE: 1999-03-03
;  EARLIER APPLICATION NUMBER: 1013/96
;  EARLIER FILING DATE: 1996-09-17
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 201
;   TYPE: PRT
;   ORGANISM: Cellulase variants
US-09-261-329-4

  Query Match             37.0%;  Score 724;  DB 9;  Length 201;
  Best Local Similarity   58.7%;  Pred. No. 1.4e-43;
  Matches  121;  Conservative   37;  Mismatches   42;  Indels    6;  Gaps    2;

Qy        154 GNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLADNNTQNGCVGGSSYTCNDNQP 213
              |:|::|||||||||||:| ||| |  || :|: :  | | | ::|| |||:||| :| |
Db          1 GSGKSTRYWDCCKPSCAWSGKASVNRPVLACDANNNPLNDANVKSGCDGGSAYTCANNSP 60

Qy        214 WVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTGSDLGSNTGA 273
              | |:|:|||||| :|||:|::|||||: ||||| | || :||| |:|| |||||
Db         61 WAVNDNLAYGFAATKLSGGTESSWCCACYALTFTSGPVSGKTLVVQSTSTGGDLGSN--- 117

Qy        274 HFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKWRFGWFK 333
              |||| |||||||:::|| ::|      ||:|||:|| |:| : |:||: ||:||| |||
Db        118 HFDLNMPGGGVGLFDGCKREFGGLP---GAQYGGISSRSECDSFPAALKPGCQWRFDWFK 174

Qy        334 NADNPTMTYKQVTCPKAITAKSGCSR 359
              |||||  |:||| ||  :|:::|| |
Db        175 NADNPEFTFKQVQCPSELTSRTGCKR 200


RESULT 8
US-09-735-787-4
; Sequence 4, Application US/09735787
; Patent No. US20010036910A1
;  GENERAL INFORMATION:
;       APPLICANT: Rasmussen, Grethe
;                  Mikkelsen, Jan Moller
;                  Schulein, Martin
;                  Patkar, Shankant A.
;                  Hagen, Fred
;       TITLE OF INVENTION: A Cellulase Preparation Comprising an
;                           Endoglucanase Enzyme
;       NUMBER OF SEQUENCES: 33
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: No. US20010036910A1o No. US20010036910A1disk of No. US200100369
;            STREET: 405 Lexington Avenue, 64th Floor
;            CITY: New York
;            STATE: New York
;            COUNTRY: United States of America
;            ZIP: 10174-6401
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/735,787
;            FILING DATE: 13-Dec-2000
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/189,028
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Lambiris, Elias J.
;            REGISTRATION NUMBER: 33,728
;            REFERENCE/DOCKET NUMBER: 3469.214-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 212-867-0123
;            TELEFAX: 212-878-9655
;   INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 376 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-735-787-4

  Query Match             36.9%;  Score 722.5;  DB 10;  Length 376;
  Best Local Similarity   59.8%;  Pred. No. 3.4e-43;
  Matches  128;  Conservative   31;  Mismatches   48;  Indels    7;  Gaps    4;

Qy        147 AVSGGASGNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLADNNTQNGCV-GGSS 205
              |||  |||:| :|||||||||||||| ||| | :|  :|:|:   ::: |  |||  |||:
Db         14 AVS-AASGSGHSTRYWDCCKPSCSWSGKAAVNAPALTCDKNDNPISNTNAVNGCEGGGSA 72

Qy        206 YTCNDNQPWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTGS 265
              | | :  || |:|:|||||||  |||||||:|||||: ||||:  |||||:|| ||||
Db         73 YACTNYSPWAVNDELAYGFAATKISGGSEASWCCACYALTFTTGPVKGKKMIVQSTNTGG 132

Qy        266 DLGSNTGAHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGC 325
```

```
Db        133 DLGDN---HFDLMMPGGGVGIFDGCTSEFGKALG--GAQYGGISSRSECDSYPELLKDGC 187

Qy        326 KWRFGWFKNADNPTMTYKQVTCPKAITAKSGCSR 359
Db        188 HWRFDWFENADNPDFTFEQVQCPKALLDISGCKR 221


RESULT 9
US-09-261-329-6
; Sequence 6, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Kim
;  APPLICANT: Schulein, Martin
;  APPLICANT: Christiansen, Lars
;  APPLICANT: Damgaard, Bo
;  APPLICANT: Von Der Osten,  Claus
;  TITLE OF INVENTION: Cellulase Variants
;  FILE REFERENCE: 4887.204-US
;  CURRENT APPLICATION NUMBER: US/09/261,329
;  CURRENT FILING DATE: 1999-03-03
;  EARLIER APPLICATION NUMBER: 1013/96
;  EARLIER FILING DATE: 1996-09-17
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
;   LENGTH: 203
;   TYPE: PRT
;   ORGANISM: Cellulase variants
US-09-261-329-6

  Query Match           36.5%;  Score 713;  DB 9;  Length 203;
  Best Local Similarity   59.4%;  Pred. No. 8.2e-43;
  Matches  123;  Conservative   31;  Mismatches   47;  Indels    6;  Gaps    3;

Qy        154 GNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLADNNTQNGCV-GGSSYTCNDNQ 212
Db          1 GSGHSTRYWDCCKPSCSWSGKAAVNAPALTCDKNDNPISNTNAVNGCEGGGSAYACTNYS 60

Qy        213 PWVVSDDLAYGFAAASISGGSEATWCCACPELTFTSTAVKGKKMVVQVTNTGSDLGSNTG 272
Db         61 PWAVNDELAYGFAATKISGGSEASWCCACYALTFTTGPVKGKKMIVQSTNTGGDLGDN-- 118

Qy        273 AHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKWRFGWF 332
Db        119 -HFDLMMPGGGVGIFDGCTSEFGKALG--GAQYGGISSRSECDSYPELLKDGCHWRFDWF 175

Qy        333 KNADNPTMTYKQVTCPKAITAKSGCSR 359
Db        176 ENADNPDFTFEQVQCPKALLDISGCKR 202


RESULT 10
US-10-007-521-10
; Sequence 10, Application US/10007521
; Publication No. US20030054539A1
;   GENERAL INFORMATION:
;        APPLICANT: Schulein, Martin
;                   Andersen, Lene N.
;                   Lassen, Soren F.
;                   Kauppinen, Markus S.
;                   Lange, Lene
;                   Nielsen, Ruby I.
;                   Ihara, Michiko
;                   Takagi, Shinobu
;        TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
;        NUMBER OF SEQUENCES: 109
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US200300545
;             STREET: 405 Lexington Avenue, 64th Floor
;             CITY: New York
;             STATE: New York
;             COUNTRY: United States of America
;             ZIP: 10174-6401
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/007,521
;             FILING DATE: 10-Dec-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/651,136
;             FILING DATE: 21-MAY-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Lambiris, Elias J.
;             REGISTRATION NUMBER: 33,728
;             REFERENCE/DOCKET NUMBER: 4366.200-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 212-867-0123
;             TELEFAX: 212-878-9655
;   INFORMATION FOR SEQ ID NO: 10:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 349 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: protein
;        SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-007-521-10

  Query Match           36.3%;  Score 710.5;  DB 9;  Length 349;
  Best Local Similarity   58.2%;  Pred. No. 2.2e-42;
  Matches  121;  Conservative   31;  Mismatches   51;  Indels    5;  Gaps    2;

Qy        152 ASGNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLADNNTQNGCVGGSSYTCNDN 211
Db         21 ASGKGHTTRYWDCCKTSCAWEGKASVSEPVLTCNKQDNPIVDANARSGCDGGGAFACTNN 80

Qy        212 QPWVVSDDLAYGFAAASISGGSEATWCCACPELTFTSTAVKGKKMVVQVTNTGSDLGSNT 271
Db         81 SPWAVSEDLAYGFAATALSGGTEGSWCCACYAITFTSGPVAGKKMVVQSTNTGGDLSNN- 139

Qy        272 GAHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKWRFGW 331
Db        140 --HFDLMIPGGGLGIFDGCSAQFGQLLP--GERYGGVSSRSQCDGMPELLKDGCQWRFDW 195

Qy        332 FKNADNPTMTYKQVTCPKAITAKSGCSR 359
Db        196 FKNSDNPDIEFEQVQCPKELIAVSGCVR 223


RESULT 11
US-10-007-521-22
; Sequence 22, Application US/10007521
; Publication No. US20030054539A1
;   GENERAL INFORMATION:
;        APPLICANT: Schulein, Martin
;                   Andersen, Lene N.
;                   Lassen, Soren F.
;                   Kauppinen, Markus S.
;                   Lange, Lene
;                   Nielsen, Ruby I.
;                   Ihara, Michiko
;                   Takagi, Shinobu
;        TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
;        NUMBER OF SEQUENCES: 109
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US200300545
;             STREET: 405 Lexington Avenue, 64th Floor
;             CITY: New York
;             STATE: New York
;             COUNTRY: United States of America
;             ZIP: 10174-6401
```

```
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/007,521
;            FILING DATE: 10-Dec-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/651,136
;            FILING DATE: 21-MAY-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Lambiris, Elias J.
;            REGISTRATION NUMBER: 33,728
;            REFERENCE/DOCKET NUMBER: 4366.200-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 212-867-0123
;            TELEFAX: 212-878-9655
;   INFORMATION FOR SEQ ID NO: 22:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 310 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-007-521-22

  Query Match           35.7%;  Score 698.5;  DB 9;  Length 310;
  Best Local Similarity   56.7%;  Pred. No. 1.3e-41;
  Matches  122;  Conservative   33;  Mismatches   53;  Indels    7;  Gaps    4;

Qy        146 SAVSGGASGNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLADNNTQNGC-VGGS 204
              : |:  :|| | ||||||||||||| |  || |: || :|:::   || :  ::||   |
Db         12 AGVAAQSSGTGRTTRYWDCCKPSCGWDEKASVSQPVKTCDRNNNPLA-STARSGCDSNGV 70

Qy        205 SYTCNDNQPWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTG 264
              :|||||||||| |:|:|||||||| : ||||||:|||||: | |||  | || |||| ||||
Db         71 AYTCNDNQPWAVNDNLAYGFAATAFSGGSEASWCCACYALQFTSGPVAGKTMVVQSTNTG 130

Qy        265 SDLGSNTGAHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAG 324
               ||  |   |||: |||||:||::||  |||       | |||| :| | || :||||| |
Db        131 GDLSGN---HFDILMPGGGLGIFDGCTPQWGVSFP--GNRYGGTTSRSQCSQIPSALQPG 185

Qy        325 CKWRFGWFKNADNPTMTYKQVTCPKAITAKSGCSR 359
              | ||: || :|||| :::::| || |:| ::|| |
Db        186 CNWRYDWFNDADNPDVSWRRVQCPAALTDRTGCRR 220


RESULT 12
US-09-261-329-3
; Sequence 3, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Kim
;  APPLICANT: Schulein, Martin
;  APPLICANT: Christiansen, Lars
;  APPLICANT: Damgaard, Bo
;  APPLICANT: Von Der Osten,  Claus
;  TITLE OF INVENTION: Cellulase Variants
;  FILE REFERENCE: 4887.204-US
;  CURRENT APPLICATION NUMBER: US/09/261,329
;  CURRENT FILING DATE: 1999-03-03
;  EARLIER APPLICATION NUMBER: 1013/96
;  EARLIER FILING DATE: 1996-09-17
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 202
;   TYPE: PRT
;   ORGANISM: cellulase variants
US-09-261-329-3

  Query Match           35.3%;  Score 689.5;  DB 9;  Length 202;
  Best Local Similarity   58.0%;  Pred. No. 3.6e-41;
  Matches  120;  Conservative   30;  Mismatches   50;  Indels    7;  Gaps    4;

Qy        154 GNGETTRYWDCCKPSCSWPGKADVTSPVGSCNKDGKTLADNNTQNGC-VGGSSYTCNDNQ 212
              | | |||||||||||| |  || |: || :|:::   || :  ::||   | :|||||||
Db          1 GTGRTTRYWDCCKPSCGWDEKASVSQPVKTCDRNNNPLA-STARSGCDSNGVAYTCNDNQ 59

Qy        213 PWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTGSDLGSNTG 272
              || |:|:||||||| : ||||||:|||||: | |||  | || |||| |||| || :|
Db         60 PWAVNDNLAYGFAATAFSGGSEASWCCACYALQFTSGPVAGKTMVVQSTNTGGDLSGN-- 117

Qy        273 AHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKWRFGWF 332
              |||: |||||:||::||  |||       | |||| :| | || :||||| || ||: ||
Db        118 -HFDILMPGGGLGIFDGCTPQWGVSFP--GNRYGGTTSRSQCSQIPSALQPGCNWRYDWF 174

Qy        333 KNADNPTMTYKQVTCPKAITAKSGCSR 359
              :|||| :::::| || |:| ::|| |
Db        175 NDADNPDVSWRRVQCPAALTDRTGCRR 201


RESULT 13
US-09-261-329-9
; Sequence 9, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Kim
;  APPLICANT: Schulein, Martin
;  APPLICANT: Christiansen, Lars
;  APPLICANT: Damgaard, Bo
;  APPLICANT: Von Der Osten,  Claus
;  TITLE OF INVENTION: Cellulase Variants
;  FILE REFERENCE: 4887.204-US
;  CURRENT APPLICATION NUMBER: US/09/261,329
;  CURRENT FILING DATE: 1999-03-03
;  EARLIER APPLICATION NUMBER: 1013/96
;  EARLIER FILING DATE: 1996-09-17
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 9
;   LENGTH: 203
;   TYPE: PRT
;   ORGANISM: Cellulase variants
US-09-261-329-9

  Query Match           35.1%;  Score 687;  DB 9;  Length 203;
  Best Local Similarity   57.0%;  Pred. No. 5.4e-41;
  Matches  118;  Conservative   35;  Mismatches   46;  Indels    8;  Gaps    4;

Qy        155 NGETTRYWDCCKPSCSWPGKADVTSPVGSC--NKDGKTLADNNTQNGCVGGSSYTCNDNQ 212
              :| |||||||||||:| ||| |: |||:|  | : :| :|   :: | |||:| |::
Db          2 SGVTTRYWDCCKPSCAWTGKASVSKPVGTCDINDNAQTPSD-LLKSSCDGGSAYYCSNQG 60

Qy        213 PWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTGSDLGSNTG 272
              || |:| |:|||||| :||  |  ||| |::|||||||| ||:|:||:|||| |||:|
Db         61 PWAVNDSLSYGFAAAKLSGKQETDWCCGCYKLTFTSTAVSGKQMIVQITNTGGDLGNN-- 118

Qy        273 AHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKWRFGWF 332
              |||: ||||||||:|||: ||       | :||| :  | |: |||  || | ||| ||
Db        119 -HFDIAMPGGGVGIFNGCSKQWNG--INLGNQYGGFTDRSQCATLPSKWQASCNWRFDWF 175

Qy        333 KNADNPTMTYKQVTCPKAITAKSGCSR 359
              :|||||: :: |||| : |::|||
Db        176 ENADNPTVDWEPVTCPQELVARTGCSR 202


RESULT 14
US-10-007-521-14
; Sequence 14, Application US/10007521
; Publication No. US20030054539A1
;   GENERAL INFORMATION:
;        APPLICANT: Schulein, Martin
;                   Andersen, Lene N.
```

```
;                 Lassen, Soren F.
;                 Kauppinen, Markus S.
;                 Lange, Lene
;                 Nielsen, Ruby I.
;                 Ihara, Michiko
;                 Takagi, Shinobu
;       TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
;       NUMBER OF SEQUENCES: 109
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US200300545
;            STREET: 405 Lexington Avenue, 64th Floor
;            CITY: New York
;            STATE: New York
;            COUNTRY: United States of America
;            ZIP: 10174-6401
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/007,521
;            FILING DATE: 10-Dec-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/651,136
;            FILING DATE: 21-MAY-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Lambiris, Elias J.
;            REGISTRATION NUMBER: 33,728
;            REFERENCE/DOCKET NUMBER: 4366.200-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 212-867-0123
;            TELEFAX: 212-878-9655
;   INFORMATION FOR SEQ ID NO: 14:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 222 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-007-521-14

  Query Match             35.1%;  Score 687;  DB 9;  Length 222;
  Best Local Similarity   57.0%;  Pred. No. 6e-41;
  Matches  118;  Conservative   35;  Mismatches   46;  Indels    8;  Gaps    4;

Qy        155 NGETTRYWDCCKPSCSWPGKADVTSPVGSC--NKDGKTLADNNTQNGCVGGSSYTCNDNQ 212
              :| |||||||||||||:| ||| |: |||:|  | : :| :|   :: | |||:| |::
Db         21 SGVTTRYWDCCKPSCAWTGKASVSKPVGTCDINDNAQTPSD-LLKSSCDGGSAYYCSNQG 79

Qy        213 PWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTGSDLGSNTG 272
              || |:| |:|||||| :||  |  ||| |::|||||||| ||:|:||:|||| |||:|
Db         80 PWAVNDSLSYGFAAAKLSGKQETDWCCGCYKLTFTSTAVSGKQMIVQITNTGGDLGNN-- 137

Qy        273 AHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKWRFGWF 332
               |||: |||||||:|||: ||       | :||| :  | |: |||  || | ||| ||
Db        138 -HFDIAMPGGGVGIFNGCSKQWNG--INLGNQYGGFTDRSQCATLPSKWQASCNWRFDWF 194

Qy        333 KNADNPTMTYKQVTCPKAITAKSGCSR 359
              :|||||: :: ||||: : |::||||
Db        195 ENADNPTVDWEPVTCPQELVARTGCSR 221


RESULT 15
US-10-007-521-24
; Sequence 24, Application US/10007521
; Publication No. US20030054539A1
;   GENERAL INFORMATION:
;       APPLICANT: Schulein, Martin
;                  Andersen, Lene N.
;                  Lassen, Soren F.
;                  Kauppinen, Markus S.
;                  Lange, Lene
;                  Nielsen, Ruby I.
;                  Ihara, Michiko
;                  Takagi, Shinobu
;       TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
;       NUMBER OF SEQUENCES: 109
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US200300545
;            STREET: 405 Lexington Avenue, 64th Floor
;            CITY: New York
;            STATE: New York
;            COUNTRY: United States of America
;            ZIP: 10174-6401
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/10/007,521
;            FILING DATE: 10-Dec-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/651,136
;            FILING DATE: 21-MAY-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Lambiris, Elias J.
;            REGISTRATION NUMBER: 33,728
;            REFERENCE/DOCKET NUMBER: 4366.200-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 212-867-0123
;            TELEFAX: 212-878-9655
;   INFORMATION FOR SEQ ID NO: 24:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 294 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-10-007-521-24

  Query Match             35.1%;  Score 687;  DB 9;  Length 294;
  Best Local Similarity   57.0%;  Pred. No. 8.1e-41;
  Matches  118;  Conservative   35;  Mismatches   46;  Indels    8;  Gaps    4;

Qy        155 NGETTRYWDCCKPSCSWPGKADVTSPVGSC--NKDGKTLADNNTQNGCVGGSSYTCNDNQ 212
              :| |||||||||||||:| ||| |: |||:|  | : :| :|   :: | |||:| |::
Db         21 SGVTTRYWDCCKPSCAWTGKASVSKPVGTCDINDNAQTPSD-LLKSSCDGGSAYYCSNQG 79

Qy        213 PWVVSDDLAYGFAAASISGGSEATWCCACFELTFTSTAVKGKKMVVQVTNTGSDLGSNTG 272
              || |:| |:|||||| :||  |  ||| |::|||||||| ||:|:||:|||| |||:|
Db         80 PWAVNDSLSYGFAAAKLSGKQETDWCCGCYKLTFTSTAVSGKQMIVQITNTGGDLGNN-- 137

Qy        273 AHFDLQMPGGGVGIYNGCATQWGAPTDGWGARYGGVSSASDCSNLPSALQAGCKWRFGWF 332
               |||: |||||||:|||: ||       | :||| :  | |: |||  || | ||| ||
Db        138 -HFDIAMPGGGVGIFNGCSKQWNG--INLGNQYGGFTDRSQCATLPSKWQASCNWRFDWF 194

Qy        333 KNADNPTMTYKQVTCPKAITAKSGCSR 359
              :|||||: :: ||||: : |::||||
Db        195 ENADNPTVDWEPVTCPQELVARTGCSR 221


Search completed: June 18, 2003, 17:44:43
Job time : 78.2272 secs
```